

{mvc,pb}@fct.unl.pt  
CENTRIA - Centre for Artificial Intelligence  
Departamento de Informática  
FCT/UNL Quinta da Torre 2829-516 CAPARICA - Portugal  
Tel. (+351) 21 294 8536  
FAX (+351) 21 294 8541

# View-based propagation of decomposable constraints

Marco Correia      Pedro Barahona

January 21, 2013

arXiv:1301.4272v1 [cs.AI] 17 Jan 2013

The final publication is available at <http://link.springer.com>

Constraints that may be obtained by composition from simpler constraints are present, in some way or another, in almost every constraint program. The decomposition of such constraints is a standard technique for obtaining an adequate propagation algorithm from a combination of propagators designed for simpler constraints. The decomposition approach is appealing in several ways. Firstly because creating a specific propagator for every constraint is clearly infeasible since the number of constraints is infinite. Secondly, because designing a propagation algorithm for complex constraints can be very challenging. Finally, reusing existing propagators allows to reduce the size of code to be developed and maintained. Traditionally, constraint solvers automatically decompose constraints into simpler ones using additional auxiliary variables and propagators, or expect the users to perform such decomposition themselves, eventually leading to the same propagation model. In this paper we explore views, an alternative way to create efficient propagators for such constraints in a modular, simple and correct way, which avoids the introduction of auxiliary variables and propagators.

## 1 Introduction

Most specialized filtering algorithms, i.e. propagators, available in constraint solvers target constraints with a specific algebraic structure. When the problem to be solved has a constraint for which no propagator exists one option is to decompose it into some logically equivalent formula involving simpler constraints. This can be done either by the user, or automatically if the solver supports it, by introducing a set of auxiliary variables and propagators. In practice the constraints which are most often decomposed are those involving multiple distinct arithmetic or logical expressions. Throughout this paper we will refer to constraints for which no specialized filtering algorithm is appropriate, and which are therefore considered for decomposition, as *decomposable constraints*.

In this paper we describe view-based propagation - a method for propagating decomposable constraints that avoids the use of auxiliary variables and propagators. Although this method does not improve the strength of the propagation algorithm, it results in significantly faster propagation compared with the decompositions obtained with auxiliary variables for the specific class of bounds completeness. An important aspect of view based propagation is that it does not preclude the introduction of auxiliary variables when required, meaning it can be combined with other modeling techniques such as subexpression substitution or variable elimination. It can also be used with specialized propagation algorithms, e.g. used in global constraints (or even replace them), allowing the exploitation of the decomposition approach combined with the efficiency of such dedicated algorithms.

The problem of propagating decomposable constraints may be approached using knowledge compilation techniques [15, 7]. These methods create a compact tractable representation of the set of solutions to the constraint and apply a general propagation algorithm which filters tuples not found in this set. However, due to its expensive runtime complex-

ity applying these methods with arithmetic expressions is only practical when expressions have small domains. In contrast, view-based propagators for decomposable constraints, extending previous proposals for propagators (e.g. indexical constraints [17, 6]), do not require exponential memory.

Rina Dechter [22] approaches decomposition of constraints from a different perspective. There, the full constraint network is taken into consideration in order to obtain generic global search algorithms with theoretical performance guarantees. The propagation is still time or space exponential but depends exclusively on specific graph-based parameters of the graph describing the constraint network. In contrast, we focus on the decomposition of a single constraint into simpler constraints for which local propagation algorithms are applied independently.

View based-propagation was introduced in [10] and [24]. The former presents a general overview of a constraint solver incorporating type polymorphism and its application for creating propagators from decomposable expressions. The latter coins the term *view* (we originally called them *polymorphic constraints*), and describes how it can be used for creating generic propagator implementations, that is, propagators which can be reused for different constraints. The present work can be seen as an extension of [26, 25] for allowing the use of a particular kind of view over functions involving multiple variables ([26] restricts the use of views to injective functions and therefore to unary functions mostly). We adapted and extended its formalization to present and prove important properties of our model. Additionally, our framework employs views as modeling primitives by automatically deriving new propagators for decomposable constraints present in a problem. In contrast, views as described in [26] are used essentially as a development tool for increasing the number of available propagators in the library.

Compilers for constraint modeling languages [21, 11, 1] generate efficient constraint solvers from a high level description of a constraint problem. The solver generation process may avoid introducing auxiliary variables for a decomposable constraint whenever a specific propagator is available. In this paper we will show that, in the absence of a specific propagator for a given constraint, a view-based propagator may be an appealing alternative.

Indexicals [17, 6] and constrained expressions [18] are conceptually close to the idea of views described in this paper. Like views, these techniques intend to support automatic propagation of decomposable constraints, however they require extra work from the user (in the case of indexicals) or the CPU (in the case of constrained expressions) when compared to our approach. These (dis)similarities will be discussed after the necessary background is presented.

This paper is organised as follows. The following section overviews the main concepts of constraint programming, introducing domains and domain approximations as well as constraint propagators and their main properties. Section 3 presents some examples of decomposable constraints and how to express propagators for such constraints, discussing the properties of various approximations, namely soundness and completeness. It also introduces box view propagators, the propagation model that will be used throughout the paper. Section 4 briefly addresses two alternatives for implementing box views in strongly

typed programming languages (exploiting either subtype or parametric polymorphism). Section 5 presents experimental results for a comprehensive set of benchmarks. The paper concludes in section 6 with a summary of the results obtained and some suggestions for future work.

## 2 Background: Domains and Propagators

This section presents the necessary concepts and notation for describing propagation in detail.

### 2.1 Constraint Satisfaction Problems

A constraint satisfaction problem (CSP) is a triple  $\langle X, D, C \rangle$  where  $X$  is a finite set of variables,  $D$  is a finite set of variable domains, and  $C$  is a finite set of constraints. We refer to the set of constraints involving some variable  $x \in X$  as  $C(x)$  and the set of variables in some constraint  $c \in C$  as  $X(c)$ .

An  $n$ -ary constraint (on variables  $x_1, \dots, x_n$ ) is a set of tuples on its  $n$  variables ( $n$ -tuples). Although tuples of integers will be most often used, we only expect that the elements in a tuple are from totally ordered types. In general, we will denote an  $n$ -tuple as  $\mathbf{x}^n$  and a set of  $n$ -tuples as  $S^n$  (or simply  $\mathbf{x}$  and  $S$  respectively when there is no ambiguity on its arity) and the set of tuples of constraint  $c$  as  $\text{con}(c)$ .

Implicitly, a constraint restricts the values of each of its variables. Denoting by  $\text{proj}_i(S^n)$  or simply  $S_i^n$  the projection of the set of tuples to their  $i$ -th element, constraint  $c$  restricts its  $i$ -th variable to take values in  $\text{con}(c)_i$ . In particular, the variable domain constraint  $D(x)$  is a special case of a unary constraint, restricting the initial values of variable  $x$ .

### 2.2 Domain Approximations

Following [26] we generalise the notion of domains from single variables to more general  $n$ -ary domains and characterize their approximations. First we denote by  $\text{conv}(S^1)$  the convex hull of some set  $S^1$  from an ordered type  $\mathcal{D}$  i.e.

$$\text{conv}_{\mathcal{D}}(S^1) = \{z \in \mathcal{D} : \min(S^1) \leq z \leq \max(S^1)\}$$

We introduce now important tuple set operations: Cartesian approximation (see for example [2]) and box approximation [4].

**Definition 2.1** (Cartesian approximation). The Cartesian approximation  $S^\delta$  (or  $\delta$ -approximation) of a tuple set  $S \subseteq \mathbb{Z}^n$  is the smallest Cartesian product which contains  $S$ , that is:

$$S^\delta = \text{proj}_1(S) \times \dots \times \text{proj}_n(S)$$

**Definition 2.2** (Box approximation). The box approximation  $S^{\beta(\mathcal{D})}$  (or  $\beta$ -approximation) of a tuple set  $S \subseteq \mathcal{D}^n$  is the smallest  $n$ -dimensional box containing  $S$ , that is:

$$S^{\beta(\mathcal{D})} = \text{conv}_{\mathcal{D}}(\text{proj}_1(S)) \times \dots \times \text{conv}_{\mathcal{D}}(\text{proj}_n(S))$$

We will address integer and real box approximations, i.e.  $S^{\beta(\mathbb{Z})}$  or  $S^{\beta(\mathbb{R})}$ , and denote these box approximations by  $S^\beta$  and  $S^\rho$ , respectively. Additionally, we will use the identity approximation in order to simplify notation.

**Definition 2.3** (Identity approximation). The identity operator transforms a tuple set in itself, i.e.  $S^{\mathbb{1}} = S$ .

**Proposition 2.4.** *Domain approximations have the following properties, for any  $n$ -tuple sets  $S, S_1, S_2 \subseteq \mathcal{D}^n$  and  $\Phi \in \{\mathbb{1}, \delta, \beta, \rho\}$ ,*

1. idempotency:  $(S^\Phi)^\Phi = S^\Phi$ ,
2. monotonicity:  $S_1 \subseteq S_2 \implies S_1^\Phi \subseteq S_2^\Phi$ , and
3. closure under intersection:  $S_1^\Phi \cap S_2^\Phi = (S_1^\Phi \cap S_2^\Phi)^\Phi$ . Note that in general it is not the case that  $S_1^\Phi \cap S_2^\Phi = (S_1 \cap S_2)^\Phi$ .

We may now define approximation domains, or  $\Phi$ -domains, as follows

**Definition 2.5** ( $\Phi$ -domain). A tuple set  $S$  is a  $\Phi$ -domain iff  $S = S^\Phi$  (for  $\Phi \in \{\mathbb{1}, \delta, \beta, \rho\}$ ).

**Definition 2.6.** A domain  $S_1$  is *stronger* than a domain  $S_2$  ( $S_2$  is *weaker* than  $S_1$ ), iff  $S_1 \subseteq S_2$ .  $S_1$  is *strictly stronger* than  $S_2$  ( $S_2$  is *strictly weaker* than  $S_1$ ) iff  $S_1 \subset S_2$ .

The following lemma (proofs for most propositions and lemmas in this paper are given in [8]) shows how the previously defined approximations are ordered for a given tuple set (or constraint).

**Lemma 2.7.** *Let  $S \subseteq \mathbb{Z}^n$  be an arbitrary tuple set. Then,*

$$S = S^{\mathbb{1}} \subseteq S^\delta \subseteq S^\beta \subseteq S^\rho$$

## 2.3 Propagation

**Definition 2.8** (Propagator). A propagator (or filter) implementing a constraint  $c \in C$  is a function  $\pi_c : \wp(\mathcal{D}^n) \rightarrow \wp(\mathcal{D}^n)$  that is contracting (i.e.  $\pi_c(S) \subseteq S$  for any tuple set  $S \subseteq \mathcal{D}$ ) and sound (it never removes tuples from the associated constraint, i.e.  $\text{con}(c) \cap S \subseteq \pi_c(S)$ ).

We will assume that propagators are monotonic, i.e.  $\pi_c(S_1) \subseteq \pi_c(S_2)$  if  $S_1 \subseteq S_2$ , although this property is not mandatory in modern constraint solvers, as shown in [26].

Two other properties of propagators are of interest: idempotency and completeness. A propagator  $\pi_c$  is idempotent iff  $\pi_c(\pi_c(S)) = \pi_c(S)$  for any tuple set  $S$ . We denote by  $\pi_c^*$  the iterated application of propagator  $\pi_c$  until a fixpoint is reached (hence  $\pi_c(\pi_c^*(S)) = \pi_c^*(S)$ ). Additionally, if propagator  $\pi_c$  is idempotent for any tuple set  $S$  we will refer to it as  $\pi_c^*$ .

The contracting condition alone sets a very loose upper bound on the output of a propagator. Many functions meet these requirements without performing any useful filtering (e.g. the identity function). In general, useful propagators aim at being complete with respect to some domain, thus achieving some consistency on the constraint associated with the propagator.

| Consistency                 | $\Phi\Psi$ -completeness |
|-----------------------------|--------------------------|
| domain                      | $\delta\delta$ -complete |
| $\text{bounds}(D)$          | $\delta\beta$ -complete  |
| range                       | $\beta\delta$ -complete  |
| $\text{bounds}(\mathbb{Z})$ | $\beta\beta$ -complete   |
| $\text{bounds}(\mathbb{R})$ | $\rho\rho$ -complete     |

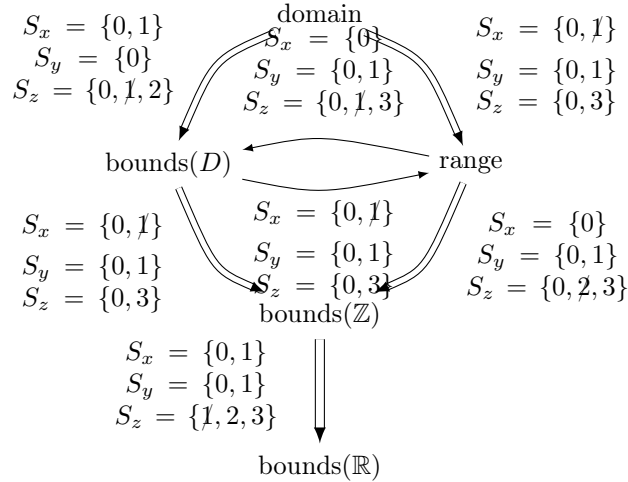
Table 1: Constraint propagator completeness.

**Definition 2.9** (Propagator completeness). A propagator  $\pi_c$  for a constraint  $c \in C$  is  $\Phi\Psi$ -complete, denoted as  $\pi_c^{\Phi\Psi}$ , iff

$$\pi_c^*(S) \subseteq (\text{con}(c) \cap S^\Phi)^\Psi$$

We should note that when considering real box approximations,  $\text{con}(c)$  in the above definition corresponds to the relaxation of constraint  $c$  to the real numbers [26]. Table 1 shows the correspondence between the constraint consistencies that are traditionally considered [26, 22, 20, 5] and the different  $\Phi\Psi$ -completeness of the propagators.

**Example 2.10** (Propagator completeness taxonomy). A hierarchy of the different types of  $\Phi\Psi$ -completenesses described is shown below, where propagators at the start of the arrows are stronger than those at their end (double arrow denotes *strictly stronger* propagators for the shown tuples). The figure considers constraint  $c = [2x + 3y = z]$ , and the arrows are labelled with tuple sets  $S = S_x \times S_y \times S_z$ , where the crossed-out values are removed by the stronger but not by the weaker propagator. For example, when applied to a domain  $S = \{\langle 0, 1 \rangle, \langle 0 \rangle, \langle 0, 1, 2 \rangle\}$ , a domain ( $\delta\delta$ -complete) propagator is able to prune tuples  $\langle 0, 0, 1 \rangle$  and  $\langle 1, 0, 1 \rangle$  that a  $\text{bounds}(D)$  ( $\delta\beta$ -complete) propagator cannot.



Notice that propagators achieving  $\text{bounds}(D)$  and  $\text{range}$  consistency are incomparable, as illustrated by the tuple sets that label the two *single* arrows with opposing directions.

Note also that  $\text{bounds}(\mathbb{R})$  completeness is weaker than  $\text{bounds}(\mathbb{Z})$  completeness, as shown in the figure. In particular, whereas a  $\text{bounds}(\mathbb{Z})$  complete propagator prunes value  $z = 1$  from the tuples in the figure, this is not the case with  $\text{bounds}(\mathbb{R})$  complete propagators since  $\langle 0.5, 0, 1 \rangle$  is a solution of the relaxation of constraint to the real numbers.



### 3 Views for Propagation of Decomposable Constraints

This section describes propagation of decomposable constraints as a function of *views*, an abstraction representing the pointwise evaluation of a function over a given (tuple) set. First we characterize decomposable constraints (3.1). Then we introduce views (3.2), and show that they may be used to express sound propagators for decomposable constraints (3.3). Section 3.4 briefly discusses propagation strength of view-based propagators, a subject that will be revisited later in section 5.3.4 after the implementation details are presented. Finally, we introduce box view-based propagators (3.5), a special case of view-based propagators that may be implemented efficiently in a straightforward way, as described in the rest of the paper.

#### 3.1 Decomposable constraints

As previously discussed, constraint solvers do not have a specific propagator for all possible constraints. Constraints that cannot be captured by a single specific propagator are usually decomposed into a logically equivalent conjunction of simpler constraints for which specific propagators exist. Many types of constraints used in practice are decomposable in this sense. Below are some examples.

**Example 3.1** (Arithmetic constraints). These are probably the most common type of decomposable constraints. Examples are constraints involving a sum of variables, e.g.  $[\sum_i x_i = k]$ , a linear combination, e.g.  $[\sum_i a_i x_i \geq k]$ , or a product, e.g.  $[\prod_i x_i = k]$ . While some solvers do have specific propagators for these constraints (e.g. by using the notion of linear relation), that approach does not work for more irregular arithmetic constraints that may include an arbitrary combination of arithmetic operators, such as for example  $[|x_1 - x_2| = 2x_3]$ .

**Example 3.2** (Boolean constraints). Boolean constraints involve Boolean domain variables or expressions, for example a disjunction of variables  $[\bigvee_i x_i]$ , or more complex expressions on Boolean constraints such as disjunctions  $[\bigvee_i (x_i > 0)]$ , logical implications  $[x > 0 \Rightarrow y < 0]$  or equivalences  $[x > 0 \Leftrightarrow y]$ .

**Example 3.3** (Counting constraints). Counting constraints restrict the number of occurrences of some values within a collection of variables, for example the EXACTLY( $\mathbf{x}, v, c$ ) constraint  $[\sum_i (x_i = v) = c]$ , or the AMONG( $\mathbf{x}, V, c$ ) constraint  $[\sum_i (x_i \in V) = c]$ .

**Example 3.4** (Data constraints). Also known as *ad-hoc* constraints, they represent an access to an element of a data structure (a table, a matrix, a relation) [3]. The most common constraint in this class is the ELEMENT( $\mathbf{x}, i, y$ ) constraint enforcing  $[x_i = y]$  where  $i$  is a variable. Decomposable constraints involving this constraint are for example  $[x_i \geq y]$  or  $[x_i - x_j \geq 0]$ .

We propose to express decomposable constraints as a composition of functions. For this purpose we will make extensive use of functions that evaluate to tuples, i.e.  $f : \mathbb{Z}^n \rightarrow \mathbb{Z}^k$  where  $k \geq 1$ , together with tuple projections, and the usual composition and Cartesian product of functions:

**Definition 3.5** (Functional composition). Functional composition is denoted by operator  $\circ$  as usual:  $(f \circ g)(x) = f(g(x))$ .

**Definition 3.6** (Cartesian product of functions). Cartesian product of functions is denoted by operator  $\times$  as follows:  $(f \times g)(x) = \langle f(x), g(x) \rangle$ . If  $x$  is a tuple we may write

$$(f \times g)(x_1, \dots, x_n) = \langle f(x_1, \dots, x_n), g(x_1, \dots, x_n) \rangle$$

We exemplify our approach to decomposition for some of the constraints given above. In the following examples let  $x$ ,  $x_i$  and  $y$  represent variables,  $\mathbf{x}$  represent a tuple of variables  $x_i$ ,  $a$  a constant, and  $p_i(\mathbf{x}) = x_i$  the projection operator.

**Example 3.7** (Equality constraint). Let constraint  $c_{\text{eq}}(\mathbf{x}) = [x_1 = x_2]$  be the binary constraint stating that variables  $x_1$  and  $x_2$  must take the same value. A unary equality constraint  $c_{\text{eqc}}(x, a) = [x = a]$  may be obtained by  $c_{\text{eqc}} = [c_{\text{eq}} \circ (p_1 \times f_a)]$ , where  $f_a(\mathbf{x}) = a$ .

**Example 3.8** (Sum constraint). Let  $f(\mathbf{x}) = x_1 + x_2$ . A sum of three variables is represented by  $f \circ (f \times p_3)$ . The generalization to a sum of  $n$  variables is defined as

$$g = f \circ (f \times p_3) \circ (f \times p_3 \times p_4) \circ \dots \circ (f \times p_3 \times p_4 \times \dots \times p_n)$$

Therefore we may decompose a constraint for a sum of  $n$  variables  $c_{\text{sum0}}(\mathbf{x}^n) = [\sum_{i=1}^n x_i = 0]$  as  $c_{\text{sum0}} = [c_{\text{eq0}} \circ g]$ , where  $c_{\text{eq0}}(\mathbf{x}) = c_{\text{eqc}}(\mathbf{x}, 0)$  defined in the previous example.

**Example 3.9** (Linear constraint). A linear constraint  $c_{\text{lin0}}(\mathbf{a}^n, \mathbf{x}^n) = [\sum_{i=1}^n a_i x_i = 0]$  may be decomposed as follows. Let  $f_1, \dots, f_n : \mathbb{Z}^n \rightarrow \mathbb{Z}$ , where  $f_i(\mathbf{x}^n) = a_i x_i$ . Then,  $c_{\text{lin0}} = [c_{\text{sum0}} \circ (f_1 \times \dots \times f_n)]$ , where  $c_{\text{sum0}}$  is defined in the previous example.

**Example 3.10** (Arbitrary arithmetic constraint). Let  $f(\mathbf{x}) = x_1 - x_2$ ,  $g(x) = |x|$ ,  $h(x) = 2x$ . The arithmetic constraint  $c(\mathbf{x}^3) = [|x_1 - x_2| = 2x_3]$  may be represented as

$$c = [c_{\text{eq}} \circ ((g \circ f \circ (p_1 \times p_2)) \times (h \circ p_3))]$$

**Example 3.11** (Exactly constraint). This constraint, represented by  $[\sum_i (x_i = v) = c]$ , may be obtained similarly to example 3.9 but where  $f_i(\mathbf{x}^n) = [x_i = v]$  and the constraint  $c_{\text{eqc}}$  is used instead of  $c_{\text{eq0}}$ .

## 3.2 Views

In most constraint solvers, domains of subexpressions occurring in constraints are not directly available to propagators, which are designed to work exclusively with variable domains. In these solvers, an offline modeling phase is responsible for obtaining an equivalent CSP where all constraints are *flattened*, that is where each subexpression appearing in a constraint is replaced by an auxiliary variable whose domain is the domain of the subexpression.

Alternatively, we present a conceptual model that considers explicit subexpression domains while abstracting on how they are computed and maintained. This will allow us to perform a theoretical analysis of the propagation on the constraint decomposition, regardless of the method used for representing the subexpression domains.

We begin by defining a view as an abstraction which captures the domain of a subexpression in a constraint.

**Definition 3.12** (View). A view over a function  $f : \mathbb{Z}^n \rightarrow \mathbb{Z}^k$  is a pair  $\varphi = \langle \varphi_f^+, \varphi_f^- \rangle$  of two functions, the *image function*  $\varphi_f^+ : \wp(\mathbb{Z}^n) \rightarrow \wp(\mathbb{Z}^k)$ , and the *object function*  $\varphi_f^- : \wp(\mathbb{Z}^k) \rightarrow \wp(\mathbb{Z}^n)$ , defined as

$$\begin{aligned}\varphi_f^+(S^n) &= \left\{ f(\mathbf{x}^n) \in \mathbb{Z}^k : \mathbf{x}^n \in S^n \right\}, \forall S \subseteq \mathbb{Z}^n \\ \varphi_f^-(S^k) &= \left\{ \mathbf{x}^n \in \mathbb{Z}^n : f(\mathbf{x}^n) \in S^k \right\}, \forall S \subseteq \mathbb{Z}^k\end{aligned}$$

A view  $\varphi_f$  is therefore defined by considering the pointwise application of  $f$  over a given set. The image function computes a set of images of  $f$ , that is the set resulting from applying  $f$  to all the points in the given set. The object function does the inverse transformation, computing the set of objects of  $f$  for which its image is in the given set.

**Example 3.13.** Consider the view over function  $f(x) = x + 1$  and the unary tuple set  $S = \{1, 2, 3\}$ :

$$\begin{aligned}\varphi_f^+(\{1, 2, 3\}) &= \{2, 3, 4\} \\ \varphi_f^-(\{2, 3, 4\}) &= \{1, 2, 3\}\end{aligned}$$

Notice some similarities of views with indexicals introduced in [17, 6] to create propagators for arithmetic constraints. An indexical roughly corresponds to the image function  $\varphi^+$  in the sense that it computes the domain of an expression. However, unlike views, indexicals do not define the inverse transformation  $\varphi^-$  and therefore are less powerful - representing a decomposable constraint using indexicals requires the additional definition of the projection of the object function for each variable in the constraint, even if this projection may be performed automatically as shown in [6].

As discussed earlier, propagating a constraint may require consulting and updating the domain of a subexpression. A view over the subexpression defines these operations:

**Example 3.14.** Consider a view  $\varphi_g$  over function  $g(x, y) = x + y$ , and constraint  $c = [x_1 + x_2 = x_3]$  on variables  $x_1, x_2, x_3$  with domains  $D(x_1) = D(x_2) = D(x_3) = \{1, 2, 3\}$ . Function  $\varphi_g^+$  may be used to compute the subexpression domain  $D(x_1 + x_2)$  from the variable domains  $D(x_1)$  and  $D(x_2)$ , while function  $\varphi_g^-$  may be used to specify the set of values for the variables  $x_1, x_2$  that satisfy constraint  $c$ .

$$\begin{aligned}\varphi_g^+(D(x_1) \times D(x_2)) &= [2 \dots 6] \\ \varphi_g^-([2 \dots 6]) &= \{ \langle x, y \rangle \in \mathbb{Z}^2 : x + y \in [2 \dots 6] \}\end{aligned}$$

Propagation consists of removing inconsistent values from the domain of the variables. In this sense, the computation of  $\varphi_f^-$ , i.e. the set of consistent assignments, must be intersected with the original domain to guarantee that the propagation is *contracting*. The following definition captures exactly that.

**Definition 3.15** (Contracting object function). Let  $S_1 \in \mathbb{Z}^n$ ,  $S_2 \in \mathbb{Z}^k$ , be two arbitrary tuple sets and  $f : \mathbb{Z}^n \rightarrow \mathbb{Z}^k$  an arbitrary function. Then,

$$\widehat{\varphi}_f(S_2, S_1) = \varphi_f^-(S_2) \cap S_1$$

**Example 3.16.** The result of evaluating and updating the domain  $D(x_1 + x_2)$ , as defined in the previous example, may be formalized using views as follows:

$$\begin{aligned} \varphi_g^+(D(x_1) \times D(x_2)) &= [2 \dots 6] \\ \widehat{\varphi}_g([2 \dots 6], D(x_1) \times D(x_2)) &= D(x_1) \times D(x_2) \end{aligned}$$

### 3.3 View-based Propagators

Propagators for decomposable constraints may be obtained by exploring the concept of views introduced in the previous section. As illustrated above, an  $n$ -ary decomposable constraint may be regarded as a special case of functions of the form  $c \circ f$  where  $f$  is a tuple-function  $f : \mathbb{Z}^n \rightarrow \mathbb{Z}^k$  and  $c$  a constraint, mapping  $k$ -tuples to the Boolean domain.

**Definition 3.17.** Let  $c$  be a  $k$ -ary constraint for which  $\pi_c$  is a propagator and  $f$  be a tuple-function  $f : \mathbb{Z}^n \rightarrow \mathbb{Z}^k$ . A view-based propagator for constraint  $c \circ f$ , denoted as  $\tilde{\pi}_{cof}$ , is obtained by

$$\tilde{\pi}_{cof}(S^n) = \widehat{\varphi}_f\left(\pi_c^{\mathbf{1}\mathbf{1}^*}\left(\varphi_f^+(S^n)\right), S^n\right)$$

**Example 3.18.** Consider the view-based propagator for the constraint  $d = [x_1 + x_2 = x_3]$ , and let us analyse the propagation achieved on the tuple set  $S_1 = \{\langle 1, 2, 3 \rangle, \langle 4, 5, 6 \rangle\}$ .

We note that the constraint  $d$  can be decomposed into a simpler constraint  $c = [x_1 = x_2]$  and function  $g = f \times p_3$ , where function  $g$  is the Cartesian product of function  $f(\mathbf{x}) = x_1 + x_2$  applied to the 2 initial elements of the tuples  $\langle x_1, x_2, x_3 \rangle$  and  $p_3$  the projection to their third element.

Function  $\varphi_g^+$  applies the addition and projection operations to the original set,  $S_2 = \varphi_g^+(S_1) = \{\langle 3, 3 \rangle, \langle 9, 6 \rangle\}$ . The resulting set  $S_2$  is filtered by propagator  $S_3 = \pi_c^{\mathbf{1}\mathbf{1}^*}(S_2) = \{\langle 3, 3 \rangle\}$ , and transformed back into  $\widehat{\varphi}_g(S_3, S_1) = \{\langle 1, 2, 3 \rangle\}$ .

### 3.4 Approximate view-based propagators

In this section we present approximate view-based propagators and relate them to the completeness classes introduced in table 1 on page 7.

**Definition 3.19** ( $\Phi\Psi$  – complete view-based propagator). A  $\Phi\Psi$  view-based propagator for a constraint  $c \circ f$  is defined as

$$\tilde{\pi}_{cof}^{\Phi\Psi}(S) = \widehat{\varphi}_f\left(\pi_c^{\mathbf{1}\mathbf{1}^*} \circ \varphi_f^+(S^\Phi), S^\Psi\right)^\Psi \cap S, \text{ where } \Phi, \Psi \in \{\mathbf{1}, \delta, \beta\}$$

Intuitively,  $\Phi\Psi$ -completeness of a view-based propagator for a decomposable constraint of the form  $c \circ f_1 \cdots \circ f_m$  is obtained by approximating the input of the image function  $\varphi_f^+$  and the output of the object function  $\widehat{\varphi}_f$ , and not approximating the remaining view functions or propagators involved. For these view-based propagators the following property can be proved [8].

**Proposition 3.20.** *A  $\Phi\Psi$  view-based propagator for a constraint  $c \circ f$  is a  $\Phi\Psi$ -complete and idempotent propagator for  $c \circ f$ .*

The use of the above proposition is rather limited since it applies only to a propagator  $\pi_c$  that is  $\mathbb{1}\mathbb{1}$ -complete and idempotent. Achieving such completeness is usually intractable in time and/or space in general, inequality constraints being a notable exception. Moreover other approximations are performed in practical view-based propagators, which are now presented.

### 3.5 Box view-based propagators

A box view-based propagator (or simply box view propagator) is a relaxation of a  $\beta\beta$ -complete view-based propagator. In addition to the  $\beta$ -approximations already presented, it

1.  $\beta$ -approximates the output of the image function  $\varphi^+$ ,
2.  $\beta$ -approximates the input of the contracting object function  $\widehat{\varphi}$ ,
3. uses a  $\beta\beta$ -complete and idempotent propagator for  $c$

To simplify its formalization, we use the following notation for special applications of  $\Phi$ -approximations to views:

$$\begin{aligned}\Phi_f^+(S) &= \left(\varphi_f^+(S^\Phi)\right)^\Phi \\ \widehat{\Phi}_f(S_1, S_2) &= \left(\widehat{\varphi}_f(S_1^\Phi, S_2^\Phi)\right)^\Phi\end{aligned}$$

We may now formalize box view-based propagators.

**Definition 3.21** (Box view-based propagator). A box view-based propagator for a constraint  $c \circ f$  is defined as

$$\pi_{c \circ f}^\square(S) = \widehat{\beta}_f\left(\pi_c^{\beta\beta\star}\left(\beta_f^+(S)\right), S\right) \cap S$$

Box view-based propagators allow very efficient implementations since  $\beta$ -domains may be stored in constant space (the domain is fully characterized by its lower and upper bound), and computing views on  $\beta$ -domains does not depend on the size of the domain for most functions.

**Example 3.22.** Let  $f(x) = |x|$  and  $S_1, S_2 \subseteq \mathbb{Z}$  two arbitrary sets of unary tuples, and assume  $\lfloor S_2 \rfloor \geq 0$ . By definition,

$$\begin{aligned}\beta_f^+(S_1) &= \left\{ |x| : x \in S_1^\beta \right\}^\beta \\ \widehat{\beta}_f(S_2, S_1) &= \left\{ x : |x| \in S_2^\beta \wedge x \in S_1^\beta \right\}^\beta\end{aligned}$$

This definition may suggest that evaluating these functions would take linear time, but in fact they may be computed in constant time assuming that finding the minimum and maximum of  $S_1, S_2$  takes constant time:

$$\begin{aligned}\beta_f^+(S_1) &= \begin{cases} \lfloor S_1 \rfloor \dots \lceil S_1 \rceil & \Leftarrow \lfloor S_1 \rfloor > 0 \\ -\lceil S_1 \rceil \dots -\lfloor S_1 \rfloor & \Leftarrow \lceil S_1 \rceil < 0 \\ 0 \dots \max(-\lfloor S_1 \rfloor, \lceil S_1 \rceil) & \text{otherwise} \end{cases} \\ \widehat{\beta}_f(S_2, S_1) &= \begin{cases} \max(\lfloor S_2 \rfloor, \lfloor S_1 \rfloor) \dots \min(\lceil S_2 \rceil, \lceil S_1 \rceil) & \Leftarrow \lfloor S_1 \rfloor > 0 \\ \max(-\lceil S_2 \rceil, \lfloor S_1 \rfloor) \dots \min(-\lfloor S_2 \rfloor, \lceil S_1 \rceil) & \Leftarrow \lceil S_1 \rceil < 0 \\ \max(-\lceil S_2 \rceil, \lfloor S_1 \rfloor) \dots \min(\lceil S_2 \rceil, \lceil S_1 \rceil) & \text{otherwise} \end{cases}\end{aligned}$$

A thorough analysis of the propagation achieved with box view propagators is complex and dependent on the constraints involved (see [8]). Here we only present a sufficient condition for a box view based propagator to be  $\text{bounds}(\mathbb{R})$  complete.

**Proposition 3.23.** *A box view-based propagator for  $c \circ f$  is always at least  $\text{bounds}(\mathbb{R})$  complete if  $f$  is continuous.*

*Proof.*

$$\pi_{cof}^\square(S) = \widehat{\beta}_f\left(\pi_c^{\beta\beta\star}\left(\beta_f^+(S)\right), S\right) \cap S \quad (\text{def. 3.21})$$

$$= \widehat{\beta}_f\left(\left(\text{con}(c) \cap \left(\beta_f^+(S)\right)^\beta\right)^\beta, S\right) \cap S \quad (\text{def.2.9})$$

Given that  $\left(\beta_f^+(S)\right)^\beta = \beta_f^+(S)$  and  $\beta_f^-(S^\beta) = \beta_f^-(S)$ ,

$$\pi_{cof}^\square(S) = \widehat{\beta}_f\left(\text{con}(c) \cap \beta_f^+(S), S\right) \cap S$$

Since  $\text{con}(c) \subseteq \text{con}_{\mathbb{R}}(c)$ ,  $S^\beta \subseteq S^\rho$  for any  $S \subseteq \mathbb{Z}^n$ , and all operators are monotonic,

$$\pi_{cof}^\square(S) \subseteq \widehat{\rho}_f\left(\text{con}_{\mathbb{R}}(c) \cap \rho_f^+(S), S\right) \cap S$$

Note that when  $f$  is continuous the following is true,

$$\begin{aligned}\rho_f^+(S) &= \left(\varphi_f^+(S^\rho)\right)^\rho = \varphi_f^+(S^\rho) \\ \widehat{\rho}_f(S_1, S_2) &= (\widehat{\varphi}_f(S_1^\rho, S_2))^\rho = (\widehat{\varphi}_f(S_1, S_2))^\rho\end{aligned}$$

Rewriting the above expression using these equivalences gives,

$$\begin{aligned}\pi_{c \circ f}^{\square}(S) &\subseteq \left( \widehat{\varphi}_f \left( \text{con}_{\mathbb{R}}(c) \cap \varphi_f^+(S^{\rho}), S \right) \right)^{\rho} \cap S \\ &\subseteq \pi_{c \circ f}^{\rho \rho^*}(S)\end{aligned}\tag{def.2.9}$$

□

So far we have been discussing view based propagators obtained from composition of views with propagators that are idempotent. Some practical propagators are non-idempotent (or have non-idempotent implementations) for efficiency reasons. Unfortunately, the strength of a view based propagator for  $c \circ f$  is influenced by the idempotency of the underlying propagator  $\pi_c$ .

**Proposition 3.24.** *A box view-based propagator for a constraint  $c \circ f$  is stronger than a corresponding view-based propagator using a non-idempotent propagator  $\pi_c$  for  $c$ . Moreover, it may be strictly stronger.*

The fact that a view-based propagator for a constraint  $c \circ f$  with an idempotent propagator for  $c$  achieves at least the same pruning as with a non-idempotent propagator for  $c$  is not surprising since  $\pi_c^*$  is always stronger than  $\pi_c$  and all other involved functions and operators are monotonic. The following example shows it can indeed achieve more pruning.

**Example 3.25.** Let  $c \circ f = [2 \cdot x_1 \cdot x_2 = x_3]$  be a decomposable constraint, where  $c = [2 \cdot x_1 = x_2]$  and  $f(x_1, x_2, x_3) = \langle x_1 \cdot x_2, x_3 \rangle$ . A box view-based propagator  $\pi_{c \circ f}^{\square}$  applied to  $D = [2 \dots 3] \times [2 \dots 3] \times [9 \dots 15]$  leaves the domains of  $x_1$  and  $x_2$  unchanged, but prunes the domain of  $x_3$  since

$$\begin{aligned}\beta_f^+(D) &= [4 \dots 9] \times [9 \dots 15] \\ \pi_c^{\beta \beta^*}([4 \dots 9] \times [9 \dots 15]) &= [5 \dots 7] \times [10 \dots 14] \\ \widehat{\beta}_f([5 \dots 7] \times [10 \dots 14], D) \cap D &= [2 \dots 3] \times [2 \dots 3] \times [10 \dots 14]\end{aligned}$$

Now consider the following non-idempotent propagator  $\pi_c^{\beta \beta}$  for the same constraint  $c$ :

$$\pi_c^{\beta \beta} = \begin{cases} D(x_3) & \leftarrow D(x_3) \cap [\lfloor D(x_1 \times x_2) \rfloor \times 2 \dots \lceil D(x_1 \times x_2) \rceil \times 2] \\ D(x_1 \times x_2) & \leftarrow D(x_1 \times x_2) \cap [\lfloor D(x_3) \rfloor / 2 \dots \lceil D(x_3) \rceil / 2] \end{cases}$$

Unlike in the previous case, the box view propagator obtained from composition with the non-idempotent propagator  $\pi_c$  would not achieve any pruning, since

$$\begin{aligned}\beta_f^+(D) &= [4 \dots 9] \times [9 \dots 15] \\ \pi_c^{\beta \beta}([4 \dots 9] \times [9 \dots 15]) &= [5 \dots 7] \times [9 \dots 15] \\ \widehat{\beta}_f([5 \dots 7] \times [9 \dots 15], D) \cap D &= [2 \dots 3] \times [2 \dots 3] \times [9 \dots 15]\end{aligned}$$

The reason for this may be explained as follows: not being idempotent, the change in the domain of  $D(x_1 \times x_2)$  is not immediately propagated back to  $D(x_3)$  but instead is approximated back to the initial value by the encapsulating view functions,  $\beta_f^+$  and  $\hat{\beta}_f$ . Therefore, the subsequent application of the box view propagator to have no effect.

The above proposition explains the differences in propagation strength between decompositions using auxiliary variables and view objects. With variable decomposition, non-idempotent propagators for sub-expressions will always act as idempotent propagators: sub-expressions are propagated independently and will be added to the propagation queue until a fix point is reached. In theory this means that decompositions using auxiliary variables may attain stronger propagation. In practice these differences do not seem to be very significant, and are compensated by other factors as the experimental results presented later suggest.



## 4 Implementation

Implementation of views for an arbitrary expression may be complex, depending on the type of approximations the view is considering. For the specific case of box view propagators, the code that implements a view over an expression  $e$  should be no more complex than the code for a propagator of  $e = z$ , where  $e$  is an expression involving variables (i.e. not other expressions) and  $z$  is a variable. For example, implementing a box view for the expressions  $e_1 + e_2$ , or  $e_1 \times e_2$ , where  $e_i$  are expressions is very similar to implementing the propagator for  $x_1 + x_2 = x_3$ , or  $x_1 \times x_2 = x_3$ , where  $x_i$  are variables. The simplicity of these implementations is in fact a key advantage of box view propagators. The following sections provide some examples of such implementations.

### 4.1 Box view propagators in strongly typed programming languages

Implementing our conceptual box view propagator in a strongly typed programming language is possible if some sort of type polymorphism support is available in the language. Most if not all popular strongly typed programming languages have built in support for subtype polymorphism, either by overloading or through the use of inheritance in the case of object oriented programming languages. In addition, parametric polymorphism has been introduced in some object oriented programming languages such as C++, Java and C#. Parametric polymorphism allows aggressive compiler optimizations, namely function code inlining (i.e. replacing function calls by the actual code), which has a significant impact on performance as we will see later.

We have implemented box view propagators in C++. Since C++ supports both subtype and parametric polymorphism, we were able to integrate the two variants of our model within the constraint solver engine, therefore obtaining a fair experimental platform.

#### 4.1.1 Subtype polymorphism

Subtype polymorphism is available in C++ through the use of inheritance. In this setting we need to define an abstract interface for box view objects:

---

```
class Box {  
    virtual int getMin()=0;  
    virtual int getMax()=0;  
    virtual bool updMin(int i)=0;  
    virtual bool updMax(int i)=0;  
};
```

---

A box view object for a specific function implements the box view object interface (for convenience the update methods return whether the operation does not result in an empty box).

**Example 4.1.** The following class defines the subtype polymorphic box view object for the addition of two box view objects.

---

```

class Add2 : Box {
    Add2(Box ax, Box ay) : x(ax),y(ay) {}
    virtual int getMin() { return x.getMin()+y.getMin(); }
    virtual int getMax() { return x.getMax()+y.getMax(); }
    virtual bool updMin(int i)
    { return x.updMin(i-y.getMax()) and y.updMin(i-x.getMax()); }
    virtual bool updMax(int i)
    { return x.updMax(i-y.getMin()) and y.updMax(i-x.getMin()); }
    Box x;
    Box y;
};

```

---

We should note that for the sake of simplicity we omit details on the efficient copy and garbage collection of box view objects. Compiling a given constraint into subtype polymorphic box view objects is straightforward since in C++ expressions are evaluated bottom-up. Below are a set of convenience functions which may be used to create subtype polymorphic view box objects for a binary addition.

---

```

Add2 add(Box x,Box y)
{ return Add2(x,y); }

Add2 operator+(Box x,Box y)
{ return Add2(x,y); }

```

---

The user may then create box view objects for arbitrary expressions in C++ using a clean syntax:

---

```

DomVar a,b,c;
a+b*c;
add(a,mul(b,c));

```

---

From what we could infer from the available documentation, *constrained expressions* introduced in ILOG Solver [18] correspond to box view objects implemented using subtype polymorphism as shown above.

#### 4.1.2 Parametric polymorphism

The fact that the C++ compiler evaluates expressions bottom-up makes the implementation of parametric polymorphic view objects slightly more complex, since they need to be compiled top-down. The solution we propose breaks the compilation algorithm in two phases. The first phase creates a syntactic representation of the expression, called a type parametric relation object, using the natural bottom-up evaluation order intrinsic in the language. Type parametric relations capture the data and the type of the objects and operations involved in the constraint. After the full constraint is compiled, we use the obtained relation object for instantiating the required view objects.

We will use templates for defining type parametric relations, since this is the language mechanism available in C++ to support type parametric polymorphism. The following

template defines generic binary relations, where “Op” is a type describing the operator, and “X” and “Y” are types of the operands.

---

```
template<class Op, class X, class Y>
class Rel2 {
    Rel2(X x, Y y) : x(x), y(y) {}
    X x;
    Y y;
};
```

---

Since any expression may be transformed to a relation object with a unique type, we can create view objects over arbitrary expressions by defining templates over relation objects.

**Example 4.2.** The following template defines the parametric polymorphic box view object for the addition of two arbitrary objects:

---

```
template<class X, class Y>
class Box<Rel2<Add, X, Y> > {
    Box(Rel2<Add, X, Y> r) : x(r.x), y(r.y) {}
    int getMin() { return x.getMin()+y.getMin(); }
    int getMax() { return x.getMax()+y.getMax(); }
    bool updMin(int i)
    { return x.updMin(i-y.getMax()) and y.updMin(i-x.getMax()); }
    bool updMax(int i)
    { return x.updMax(i-y.getMin()) and y.updMax(i-x.getMin()); }
    Box<X> x;
    Box<Y> y;
};
```

---

Parametric relation objects are created by a set of convenience functions, such as:

---

```
template<class X, class Y>
Rel2<Add, X, Y> add(X x, Y y)
{ return Rel2<Add, X, Y>(x, y); }

template<class X, class Y>
Rel2<Add, X, Y> operator+(X x, Y y)
{ return Rel2<Add, X, Y>(x, y); }
```

---

Note that the above functions only create the relation object for the expression, and not the corresponding view object. Creating the view object is accomplished by providing the relation object to the following function (which is basically a convenience function to avoid specifying the parameter T):

---

```
template<class T>
Box<T> box(T t)
{ return Box<T>(t); }
```

---

The following code instantiates two parametric box view objects for an expression using the above constructs.

---

```
DomVar a, b, c;
box(a+b*c);
box(add(a, mul(b, c)));
```

---

## 4.2 Incrementality

It is important to note that incremental propagators (i.e. propagators which maintain a state) may be used transparently with views. This may be very useful in practice, for example to model the bounds complete distinct propagator over expressions, such as the main constraint of the Golomb ruler problem (given in the next section),

$$\text{DISTINCT}(\{x_i - x_j : 1 \leq j < i \leq m\})$$

There is also nothing preventing us from creating views that maintain an internal state. Although we have implemented this for some expressions, namely the `ELEMENT` expression, it does not seem to be useful for most  $\beta\beta$  views that are cheap to evaluate. However, it could certainly make a difference if using  $\delta\delta$  views for domain propagation. In summary, using views does not constrain the propagator implementation model to be either incremental or non-incremental.

## 4.3 Triggering

Triggering is a well known method for decreasing the number of redundant propagations during a fixpoint computation [23]. A trigger may be seen as a condition for propagation - propagators are known to be idempotent for the current domain until the condition is true, i.e. when they may perform propagation. The way triggering is used with views is not much different from its use in the variable decomposition approach. In addition to the methods presented above, each view object must provide methods for creating/deleting triggers on the relevant events. This must be adapted to the events available in the solver: for example, a trigger on both bounds of an expression  $x + y$  should map to the four bounds of  $x$  and  $y$ , whereas a trigger on the minimum of  $-x$  maps to the maximum of  $x$ .

Given that each view object provide methods for creating and removing triggers, moving triggers is also possible using views. Consider for example a box view over the expression `IFTHENELSE(C, T, F)` which represents the box  $T$  if  $C$  is true, and  $F$  if it is false. Initially the box view object for this expression creates triggers on  $C$ ,  $T$ , and  $F$  (note that inference can be made even if  $C$  is not ground). When  $C$  is instantiated it removes the trigger on either  $T$  or  $F$ . We should make two important remarks about creating/deleting triggers. Firstly, at least in our current implementation, moving triggers is not as efficient as in e.g. [14]. This is due to the fact that it is recursive on the structure of the expression (consider for example creating a trigger on an expression containing deeply nested expressions). This is tantamount with other aspects of views - a specialized propagator for the constraint will always have better runtime performance. Secondly, view objects must always check beforehand if it is safe to move a trigger. In

our example, the box view object must check if  $C$  is ground (i.e. if `getMin()==getMax()`) before deleting the relevant triggers, even if it has updated the domain of  $C$  itself. This has to do with (non) persistence of update operations explained in the following section.

#### 4.4 Persistent operations and idempotency checking

An important feature of box view objects is that operations are not guaranteed to be persistent. Consider for example the constraint  $c = [X \neq k]$ , where  $X$  is an expression and  $k$  is a constant. The pseudocode for a bounds consistent propagator for this constraint is as follows:

---

```

1 bool propagate(Box<int> X, int k) {
2     if (X.min()==k)
3         return X.updateMin(k+1);
4     if (X.max()==k)
5         return X.updateMax(k-1);
6     return true;
7 }
```

---

This propagator may be used for propagating several constraints by specifying different expressions for  $X$ , for instance the constraints  $c_1 = [x_1 \neq k]$  and  $c_2 = [y_1 + y_2 \neq k]$  where  $X$  is respectively  $x_1$  and  $y_1 + y_2$ . While lines 3 and 5 guarantee that the domain of expression  $X$  will be different from  $k$  when propagating constraint  $c_1$ , this is not guaranteed when propagating constraint  $c_2$  (consider for example  $D(y_1) = D(y_2) = \{1, 2\}$  and  $k = 4$ ). In this simple propagator the fact that some update operations are non-persistent is not important: the propagator will be scheduled again when either the minimum or maximum of  $X$  changes, eventually leading to a persistent update. In general, view based propagators must be designed to take non-persistent update operations into consideration. The class of propagators that check for idempotency deserves special attention.

Idempotency checking is a technique for decreasing the number of redundant propagations during solving. The main idea is to avoid scheduling propagators that report idempotency for a given domain since no further pruning would be achieved. A naive (but incorrect) implementation of this optimization on the propagator above would be as follows:

---

```

status propagate(Box<int> X, int k) {
    if (X.min()==k)
        return (X.updateMin(k+1))?idempotent:failed;
    if (X.max()==k)
        return (X.updateMax(k-1))?idempotent:failed;
    return suspend;
}
```

---

In the above pseudocode a propagator can be in one of three states: failed (i.e. inconsistent for the current domain), idempotent, or suspended (i.e. neither failed nor

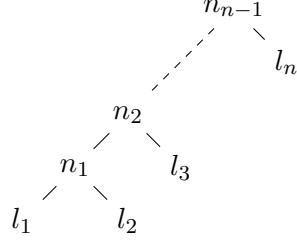


Figure 1: An unbalanced expression syntax tree. The internal nodes  $n_1 \dots n_{n-1}$  represent operators and leafs  $l_1 \dots l_n$  represent variables.

idempotent). Due to non-persistent operations for some instantiations of  $X$  this propagator may report “idempotent” when in fact it is not. Instead, a correct implementation of this optimization must not rely on persistent update operations:

---

```

status propagate(Box<int> X, int k) {
    if (X.min()==k)
    {
        if (X.updateMin(k+1))
            return (X.min()>k)?idempotent:suspend;
        else
            return failed;
    }
    if (X.max()==k)
        (similar)
    return suspend;
}

```

---

In summary, checking for idempotent view-based propagators is possible but must be carefully designed in order not to rely on persistent update operations as shown above.

## 4.5 Complexity analysis

The adoption of views avoids introducing auxiliary variables and propagators for every subexpression. Conceptually, a view object over an expression serves the same purpose as the auxiliary variable introduced for that expression: to expose its domain. However these models have distinct operational tradeoffs. To illustrate this let us focus on an arithmetic constraint involving  $n$  variables with uniform domain size  $d$ , with an unbalanced syntax tree, i.e. where each operator in the expression involves at least one variable. Figure 1 shows a fragment of the expression syntax tree. We will refer to the decomposition model using auxiliary variables as VARS, subtype polymorphic view as SVEWS, and parametric polymorphic views as PVEWS.

### 4.5.1 Memory

A box view object can be designed to expose just the subset of the expression's domain required for the view's client (e.g. the bounds of the expression). In contrast, a variable maintains the domain of the expression, possibly containing regions that will always be ignored for propagation. For an expression containing  $n-1$  operators (fig. 1), the memory overhead of the VARS model is in  $O(nD)$ , where  $D$  is the size of the largest domain of an auxiliary variable. In practice, although  $D$  may be as large as  $d^{n-1}$ , many solvers [12, 13, 18] use intervals to store the domains of the auxiliary variable (i.e.  $D = 2$ ), thus eliminating this problem.

### 4.5.2 Runtime

The analysis of runtime complexity focuses on the number of propagator executions, function calls, arithmetic operations performed and updates of variable bounds, for a single propagation of the expression. We consider two worst-case situations: a) when propagation (up to the root) is due to a change in the domain of a leaf variable, and b) when propagation to leaf variables is caused by a change in the domain of the root. These situations correspond, respectively, to accessing and updating the bounds of the expression.

- a) In the VARS model  $O(n)$  propagators may execute, forcing changes in the bounds of  $O(n)$  (auxiliary) variables with  $O(n)$  operations. In both view models, an update of a leaf variable domain causes one single propagator to execute and evaluate the full expression. Such evaluation requires  $O(n)$  function calls and  $O(n)$  operations<sup>1</sup>.
- b) In the VARS model  $O(n)$  propagators may execute, forcing changes in the bounds of up to  $O(2n)$  variables ( $n$  leaf variables and  $n-2$  auxiliary variables), requiring up to  $O(n)$  operations. Both view models force  $O(n)$  updates of bounds of the (leaf) variables, through  $O(n)$  function calls and  $O(n)$  operations, but no extra propagators will execute.

Although all the models present the same worst-case propagation complexity of  $O(n^2d)$ , breaking down the costs shows that the view models perform fewer calls to propagators and fewer variable updates (there are no auxiliary variable) than the VARS model. The fundamental operational difference between the VARS model and both view models is that a view object computes its domain on demand, that is, it will never update its domain before needed by the view's client, in contrast with the VARS model, where additional propagators will be posted to prune the domains of the auxiliary variables. Among all contributions, managing the extra propagators should involve the most significant

---

<sup>1</sup>A single update of the expression requires the evaluation of the full tree, possibly involving  $n^2$  total operations (as well as  $n^2$  function calls in case of the SViews model), if subexpressions are evaluated multiple times recursively. It is possible to decrease this number to  $n$  if the evaluation of the subexpressions is cached during the execution of the propagator. This may be done efficiently with box views since the cached data is simply the two bounds of the subexpression that do not have to be backtrackable.

costs. On the other hand, the view models require extra operations and function calls. Overall, and despite the same worst-case complexity we expect that the VARS models, that may suffer from the execution of more propagators, are out-performed by the VIEWS models, specially by the PIEWS where the compiler is often able to avoid the overhead of function calls using inlining. But these expectations can only be adequately assessed by experimentation, so in the next section we present results obtained in a comprehensive set of experiments.



## 5 Experimental Results

In this section we evaluate the performance of the decomposition methods described in the previous sections on a set of benchmarks.

### 5.1 Experiments

Specifically we are interested in comparing the following models.

#### 5.1.1 Models

**VARs.** This is the classical method for decomposing constraints into primitive propagators introducing one auxiliary variable for each subexpression.

**VARs+GLOBAL.** This model is similar to the previous but uses global constraints for lowering the number of auxiliary variables. Only a subset of problems support this decomposition in which case we will specifically mention which global constraints are used.

**PVIEWS.** The model that implements the decomposition based on parametric polymorphic view objects.

**SVIEWS.** The decomposition based on subtype polymorphic view objects.

**VIEWS+GLOBAL.** Like the VARs+GLOBAL model, this model uses a combination of some type of views and a global constraint propagator.

All the above decomposition models were implemented in CaSPER [9, 10]. Additionally, we also implemented the first two in Gecode [12] denoted GECODE-VARS and GECODE-VARS+GLOBAL respectively. Comparing to the Gecode solver assesses the competitiveness of CaSPER as a whole, in order to clarify that the performance of using our method for propagation compared to using auxiliary variables is not due to an inefficient implementation of the latter.

#### 5.1.2 Problems

The set of benchmarks covers a total of 22 instances from 6 different problems. Before we present them in detail, we should make a few general considerations.

For each given instance we used the same labeling heuristics (or no heuristics at all) for testing the above models. This means that, for each instance, the solvers resulting from the implementation of the above models explore exactly the same search space, unless the decompositions have different propagation strength, which may occur as we have already seen.

In the following, we will mostly focus on the decomposable constraints for which our models apply. When describing the model, we may choose to ignore other necessary, redundant, or symmetry breaking constraints that we used in our implementations. These

were kept constant across all implementations of the above models for each benchmark and therefore do not influence our conclusions. For additional information, we provide references to detailed descriptions of the problems in the online constraint programming benchmark database CSPLib [16]. The source code for all the implementations can be obtained from the first author upon request.

**Systems of linear equations.** This experiment consists in solving a system of linear equations. Linear equations are usually integrated in Constraint Programming using a specific global constraint propagator. The goal of the experiment is therefore to assess the overhead of decomposing expressions using the presented models compared to a decomposition which uses a special purpose algorithm, i.e. a global constraint.

Each system of linear equations is described by a tuple  $\langle n, d, c, a, s|u \rangle$  where  $n$  is the number of variables in the problem,  $d$  is the uniform domain size,  $c$  is the number of linear equations,  $a$  is the number of terms in each equation, and the last term denotes if the problem is (s)atisfiable or (u)nsatisfiable. Each problem is defined by

$$\bigwedge_{i=1}^c \sum_{v \in p(i)} v = t$$

where  $p(i)$  is a function returning a combination of  $a$  variables for the equation  $i$ , selected randomly from the full set of  $C_a^n$  possible combinations. The independent term  $t$  in each equation was selected randomly with a uniform probability from the interval  $[a \dots a \times d]$ . Different random seeds were experimented in order to generate difficult instances.

The  $a$ -ary sum constraints were decomposed using binary sums implementing a subset of the previously described models, namely VARS, SVIEWS, and PVIEWES. Model VARS+GLOBAL used global constraint propagators for the  $a$ -ary sums (in this case no auxiliary variables are required).

**Systems of nonlinear equations.** The second experiment considers systems of nonlinear equations. These problems arise often in practice, and since the decomposition to special purpose propagators is not so direct as in the previous case, it provides a realistic opportunity to apply the previously discussed models.

A system of nonlinear equations is described by a tuple  $\langle n, d, c, a_1, a_2, s|u \rangle$  where  $a_1$  is the number of terms in each equation, each term is composed of a product of  $a_2$  factors, and all remaining variables have the same meaning as before. Each system of nonlinear equations is formally defined as:

$$\bigwedge_{i=1}^c \sum_{j=1}^{a_1} \prod_{v \in p(i,j)} v = t$$

where  $p(i, j)$  is a function returning a combination of  $a_2$  variables for the term  $j$  of equation  $i$ , selected randomly from the full set of  $C_{a_2}^n$  possible combinations.

The tested models consist of the decomposition into binary sums and products using auxiliary variables exclusively (VARS) and using view models (SVIEWS, PVIEWES). We

also tested two models where each product is decomposed using either auxiliary variables or views, projected to a variable  $x_i$ , and a sum propagator is used to enforce  $\sum_{i=1}^{a_1} x_i$  for each equation (VARS+GLOBAL and PVIEW+GLOBAL respectively).

**Social golfers (prob10 in CSPLib).** The Social golfers problem consists in scheduling a golf tournament. The golf tournament lasts for a specified number of weeks  $w$ , organizing  $g$  games each week, each game involving  $s$  players. There is therefore a total of  $g \times s$  players participating in the tournament. The goal is to come up with a schedule where each pair of golfers plays in the same group at most once.

This problem may be solved efficiently in Constraint Programming using a 3-dimensional matrix  $x$  of  $w \times g \times s$  integer domain variables, where each variable identifies a golf player in the tournament. For two groups of players  $G_1, G_2$ , the MEETONCE constraint ensures that any pair of players in one group does not meet in the other group,

$$\text{MEETONCE}(G_1, G_2) = \left[ \sum_{x \in G_1, y \in G_2} (x = y) \leq 1 \right]$$

This constraint is then used to impose that each pair of players meets at most once during the entire tournament,

$$\bigwedge_{1 \leq w_i < w_j \leq w} \text{MEETONCE} \left( \left\{ \begin{array}{l} \{x_{w_i, g_i, s_i} : 1 \leq g_i \leq g, 1 \leq s_i \leq s\}, \\ \{x_{w_j, g_j, s_j} : 1 \leq g_j \leq g, 1 \leq s_j \leq s\} \end{array} \right\} \right)$$

We tested a subset of our models for propagating the MEETONCE constraint, namely PVIEW+GLOBAL, SVIEW+GLOBAL, and VARS+GLOBAL. View based models implement the MEETONCE constraint using the above expression directly using a global propagator for the sum constraint:

$$\text{MEETONCE}(G_1, G_2) = [\text{SUM}(\text{ALL}(x \in G_1, y \in G_2, x = y)) \leq 1]$$

The ALL predicate is used for aggregating a set of expressions from the instantiation of a template expression (e.g.  $x = y$ ) over a set of possible values (e.g.  $x \in G_1, y \in G_2$ ).

The VARS+GLOBAL model implements the traditional decomposition using a set  $b$  of  $s^2$  auxiliary boolean domain variables,

$$\text{MEETONCE}(G_1, G_2) = \bigwedge_{x \in G_1, y \in G_2} [b_i = (x = y)] \quad (1)$$

$$\wedge \left[ \sum_{i=1}^{s^2} b_i \leq 1 \right] \quad (2)$$

In this case we used the (reified) equality propagator for each equation in the conjunction of eq. 1, and a sum global propagator for eq. 2.

**Golomb ruler (prob6 in CSPLib).** A Golomb ruler of  $m$  marks and length  $x_m$  is a set of  $m$  integers,

$$0 = x_1 < x_2 < \dots < x_m$$

such that the  $m(m-1)/2$  differences  $x_i - x_j$ ,  $1 \leq j < i \leq m$  are distinct. The Golomb ruler problem is an optimization problem, where the goal is to find the smallest possible Golomb ruler with a given number of marks.

This problem makes use of a constraint of the form

$$\text{DISTINCT}(\{x_i - x_j : 1 \leq j < i \leq m\})$$

enforcing that the pairwise differences  $x_i - x_j$  are all distinct. The classical decomposition for this constraint (VARS+GLOBAL) introduces one auxiliary variable for each pairwise difference, and makes use of the distinct global propagator for the DISTINCT constraint,

$$\bigwedge_{1 \leq i < j \leq m} [a_{i,j} = x_i - x_j] \\ \wedge [\text{DISTINCT}(a)]$$

Using views avoids introducing the set of auxiliary variables  $a$ . Instead, the constraint is used directly as follows,

$$\text{DISTINCT}(\text{ALL}(1 \leq i \leq m, 1 \leq j \leq m, j < i, x_i - x_j))$$

and enforced with the bounds complete propagator introduced by [19]. In our benchmarks we solved the decision version of this problem, i.e. we provided the size of the ruler  $x_m$  as a parameter of the problem, and asked for a ruler satisfying the constraints.

**Low autocorrelation binary sequences (prob5 in CSPLib).** The goal is to construct a binary sequence  $S = x_1, \dots, x_n$  of length  $n$ , where  $D(x_i) = \{-1, 1\}$ ,  $1 \leq i \leq n$ , that minimizes the autocorrelations between bits, i.e. that minimizes the following expression,

$$m = \sum_{i=1}^{n-1} \left( \sum_{j=1}^{n-i-1} x_j \times x_{j+i+1} \right)^2 \quad (3)$$

The VARS+GLOBAL model decomposes the above expression using three sets of auxiliary variables  $a$ ,  $b$ , and  $c$ , and sum constraints as follows,

$$\bigwedge_{i=1}^{n-1} \bigwedge_{j=1}^{n-i-1} [x_j \times x_{j+i+1} = a_{i,j}] \\ \bigwedge_{i=1}^{n-1} \left[ \sum_{j=1}^{n-i-1} a_{i,j} = b_i \right] \\ \bigwedge_{i=1}^{n-1} [b_i^2 = c_i] \\ \wedge \left[ m = \sum_{i=1}^{n-1} c_i \right]$$

The PViews, and SViews models implements expression 3 directly, without introducing any auxiliary variable.

**Fixed-length error correcting codes (prob36 in CSPLib).** This problem involves generating a set of strings from a given alphabet which satisfy a pairwise minimum distance. Each instance is defined by a tuple  $\langle a, n, l, d \rangle$  where  $a$  is the alphabet size,  $n$  is the number of strings,  $l$  is the string length, and  $d$  is the minimum distance allowed between any two strings. For measuring the distance between two strings we have used the Hamming distance on two instances and the Lee distance on the other two. For two arbitrary strings  $x, y$  of length  $l$ , these measures are defined as follows,

$$\begin{aligned} \text{HAMMING}(x, y) &= \sum_{i=1}^l (x_i \neq y_i) \\ \text{LEE}(x, y) &= \sum_{i=1}^l \min(|x_i - y_i|, a - |x_i - y_i|) \end{aligned} \quad (4)$$

This problem was modeled by a matrix  $x$  of  $n \times l$  integer domain variables, where each variable  $x_{i,j}$  can take a value in  $1 \dots a$  corresponding to the symbol of string  $i$  in position  $j$ . Then, distance constraints are imposed between each pair of strings,

$$\bigwedge_{1 \leq i_1 < i_2 \leq n} \text{DISTANCE}(\{x_{i_1,j} : 1 \leq j \leq l\}, \{x_{i_2,j} : 1 \leq j \leq l\}) \geq d$$

The VARS+GLOBAL model decomposes distance constraints using auxiliary variables and sum constraints. Note that in the case of the Lee distance, a total of  $4l$  auxiliary variables are introduced for each distance constraint. The SVIEWS, and PVIEW models implement both distance functions without any extra variables.

## 5.2 Setup

The experiments were compiled with the gcc-4.5.3 C++ compiler and executed on an Intel Core i7 @ 3.39GHz running Mac OS X 10.7.4. The versions of the CaSPER and Gecode solvers were the most recently available at the time of these experiments, respectively version 1.0.0rc2 and version 3.7.3. Each benchmark was repeated ten times and then kept repeating until the standard deviation of the runtime was below 2% of the average time. The minimum runtime was then used. The source code for all experiments can be made available upon request (please email the first author).

## 5.3 Discussion

The results of all benchmarks are summarized in tables 2 to 6 (more detailed in [8]) from which we may draw the following conclusions.

### 5.3.1 Type parametric views versus Subtype polymorphic views

Recall from section 4.1 on page 17 that solvers using type parametric view objects are able to avoid a number of function calls due to code inlining optimizations. Table 2

|                                      | mean | stddev | min  | max  |
|--------------------------------------|------|--------|------|------|
| Systems of linear equations          | 0.44 | 1.87   | 0.24 | 0.97 |
| Systems of nonlinear equations       | 0.74 | 1.02   | 0.72 | 0.75 |
| Social golfers                       | 0.70 | 1.03   | 0.68 | 0.72 |
| Golomb ruler                         | 0.98 | 1.02   | 0.95 | 1.00 |
| Low autocorrelation binary sequences | 0.92 | 1.01   | 0.91 | 0.92 |
| Fixed-length error correcting codes  | 0.56 | 1.44   | 0.39 | 0.77 |
| All                                  | 0.67 | 1.48   | 0.24 | 1.00 |

Table 2: Geometric mean, standard deviation, minimum and maximum of the ratios between the runtimes of the solver implementing the PViews model and the solver implementing the SViews model, on all benchmarks.

shows how this optimization improves performance in practice. Overall the speed-up of PViews wrt. SViews is 50% (i.e.  $1/0.67 - 1$ ). In particular for problems involving a large number of subexpressions the speed-up can reach up to 300%. Given that type parametric views are consistently better, we choose to use this model exclusively on the remaining experiments.

### 5.3.2 Auxiliary variables versus Type parametric views

Table 3 compares the runtime of the best model using auxiliary variables, i.e. either VARS or VARS+GLOBAL, with the runtime of the best model that uses type parametric views, i.e. either PViews or PViews+GLOBAL. View objects do not intend to be a replacement for global constraint propagators, and therefore this table shows how much the runtime of a constraint program may be improved when using the best available tools.

Before we take a global view on the results in this table, let us focus on the special case of the benchmark involving systems of linear equations. We recall that this benchmark should not be considered as part of a realistic application of views or auxiliary variables since it may be modeled using a global constraint propagator exclusively. However, modeling the global sum constraint using type parametric views over binary sums was only 10% worse on average, which is nevertheless remarkable.

For all other benchmarks, using type parametric views instead of auxiliary variables was consistently better, approximately twice as fast on (geometric) average, and always more than 33% faster. An interesting particular case are the two instances of the “Fixed length error correcting codes” problem using the Lee distance. These instances are in fact the only ones for which decomposing using auxiliary variables could be recommended. This is because there is a subexpression which occurs twice in the expression, and therefore can be represented by the same auxiliary variable (see equation 4), possibly leading to a smaller search tree. Even without this optimization, the solver using type parametric views was almost twice as fast on average on these instances.

Regarding propagation, even if using auxiliary variables may sometimes lead to smaller

|                                      | mean | stddev | min  | max  |
|--------------------------------------|------|--------|------|------|
| Systems of linear equations          | 1.10 | 1.14   | 0.96 | 1.24 |
| Systems of nonlinear equations       | 0.38 | 1.08   | 0.34 | 0.41 |
| Social golfers                       | 0.28 | 1.29   | 0.22 | 0.36 |
| Golomb ruler                         | 0.75 | 1.07   | 0.70 | 0.80 |
| Low autocorrelation binary sequences | 0.24 | 1.01   | 0.24 | 0.24 |
| Fixed-length error correcting codes  | 0.60 | 1.30   | 0.42 | 0.79 |
| All                                  | 0.51 | 1.73   | 0.22 | 1.24 |
| All except linear                    | 0.43 | 1.55   | 0.22 | 0.80 |

Table 3: Geometric mean, standard deviation, minimum and maximum of the ratios between the runtimes of best performing model using views and the best performing model using auxiliary variables, on all benchmarks.

|                                     | mean | stddev | min  | max  |
|-------------------------------------|------|--------|------|------|
| Systems of nonlinear equations      | 1.06 | 1.08   | 1.01 | 1.16 |
| Golomb ruler                        | 1.00 | 1.00   | 1.00 | 1.01 |
| Fixed-length error correcting codes | 1.20 | 1.00   | 1.20 | 1.20 |
| All                                 | 1.06 | 1.08   | 1.00 | 1.20 |

Table 4: Geometric mean, standard deviation, minimum and maximum of the ratios between the number of fails of the best performing solver using views and the best performing solver using auxiliary variables, on all instances of each problem where the number of fails differ.

search trees (due to proposition 3.24), the difference was not significant in our experiments. In fact, for those instances where using auxiliary variables increases propagation strength compared to views, the discrepancy in the number of fails was only of 6% on average, and never more than 20% (table 4).

### 5.3.3 Competitiveness

Modeling decomposable constraints using type parametric views makes CaSPER competitive with the state-of-the-art Gecode solver. This may be seen by comparing the results presented in table 5 and table 6. In the first table we compare the runtimes obtained by running the same model on both solvers, i.e. VARS+GLOBAL and GECODE-VARS+GLOBAL. The second table compares the PVIEW model against GECODE-VARS+GLOBAL, which are the best models that can be implemented in both platforms using the available modeling primitives. While CaSPER is worse in all but one problem when using auxiliary variables and global propagators, it becomes faster when using type parametric views. We believe that the discrepancy observed in the “Fixed length error correcting codes” benchmark is related to aspects of the architecture of both solvers

|  | mean | stddev | min  | max  |
|--|------|--------|------|------|
| Systems of linear equations                    | 1.23 | 1.41   | 0.73 | 1.50 |
| Systems of nonlinear equations                 | 2.06 | 1.06   | 1.90 | 2.16 |
| Social golfers                                 | 2.00 | 1.28   | 1.51 | 2.38 |
| Golomb ruler                                   | 1.68 | 1.29   | 1.25 | 2.01 |
| Low autocorrelation binary sequences           | 1.84 | 1.01   | 1.83 | 1.85 |
| Fixed-length error correcting codes            | 0.11 | 2.23   | 0.04 | 0.21 |
| All  | 1.01 | 3.30   | 0.04 | 2.38 |
| All except fixed-length error correcting codes | 1.72 | 1.33   | 0.73 | 2.38 |

Table 5: Geometric mean, standard deviation, minimum and maximum of the ratios between the runtimes of the CaSPER solver implementing the VARS+GLOBAL model and the Gecode solver implementing the VARS+GLOBAL model, on all benchmarks.

|  | mean | stddev | min  | max  |
|--|------|--------|------|------|
| Systems of linear equations                    | 1.35 | 1.56   | 0.71 | 1.86 |
| Systems of nonlinear equations                 | 0.78 | 1.08   | 0.72 | 0.86 |
| Social golfers                                 | 0.57 | 1.42   | 0.45 | 0.85 |
| Golomb ruler                                   | 1.26 | 1.24   | 1.00 | 1.53 |
| Low autocorrelation binary sequences           | 0.44 | 1.02   | 0.44 | 0.45 |
| Fixed-length error correcting codes            | 0.06 | 2.26   | 0.03 | 0.13 |
| All  | 0.52 | 3.23   | 0.03 | 1.86 |
| All except fixed-length error correcting codes | 0.85 | 1.61   | 0.44 | 1.86 |

Table 6: Geometric mean, standard deviation, minimum and maximum of the ratios between the runtimes of the CaSPER solver implementing the PViews model and the Gecode solver implementing the VARS+GLOBAL model, on all benchmarks.

which are orthogonal to the tested models.

In summary, we have seen how box view objects may be implemented using several language paradigms, with a focus on strongly typed languages, namely C++. Decomposition models based on auxiliary variables, subtype polymorphic views, and type parametric views were implemented for a number of well known benchmarks, and the results were discussed. We observed that type parametric views are clearly more efficient than models resulting from the other decomposition/compilation methods for all benchmarks. Moreover, we have seen that this technique improves the performance of CaSPER to the point of being competitive with Gecode, which is regarded as one of the most efficient solvers available.



| Non-Linear            |        | p      | t    | f       | u      | calls  | op     |
|-----------------------|--------|--------|------|---------|--------|--------|--------|
| 20-<br>20-<br>10-4-   | VARs   | 1.9E+8 | 15.7 | 651,189 | 1.2E+8 | 0.0E+0 | 1.1E+9 |
|                       | SVIEWS | 2.4E+7 | 8.8  | 651,189 | 2.0E+7 | 1.5E+9 | 1.4E+9 |
|                       | PVIEWS | 2.4E+7 | 6.4  | 651,189 | 2.0E+7 | 9.7E+8 | 1.4E+9 |
| 20-5-<br>20-4-<br>3-u | VARs   | 1.6E+8 | 19.3 | 342,311 | 1.0E+8 | 0.0E+0 | 9.6E+8 |
|                       | SVIEWS | 1.2E+7 | 8.9  | 346,762 | 1.0E+7 | 1.5E+9 | 1.6E+9 |
|                       | PVIEWS | 1.2E+7 | 6.5  | 346,762 | 1.0E+7 | 5.2E+8 | 1.6E+9 |

Table 7: Number of propagator executions (p), seconds (t), fails (f), domain updates (u), calls to view object methods (calls), and arithmetic operations (op) when solving two instances of the systems of nonlinear equations problem.

### 5.3.4 Monitoring Execution

The theoretical discussion in the previous section hints at the reasons why the view models, in particular the PVIEWS model, achieve better performance than the VARs model. We checked whether these hints were confirmed in the experiments described earlier. To do so, we monitored the execution of many instances of the problems and consistently obtained results similar to those reported below. Table 7 shows the results obtained with two instances of the problem involving systems of nonlinear equations (see 5.1.2). First, we note that in the first (satisfiable) instance, the same search trees were explored by all models, whereas in the second (unsatisfiable) the VARs model explores a slightly smaller tree (1% less failures). As expected, the better performance of both view models (about 2 times faster) is due to the lower number of propagator executions as well as domain updates (one order of magnitude less than in the VARs model). Moreover, PVIEWS improves (20-30 %) on the SVIEWS model because of its better inlining, although the compiler is not able to inline all the function calls.

A similar picture is obtained with three instances of the Golomb problem presented in table 8. Again, despite exploring a slightly smaller search space (<1% less failures), the performance of the VARs model is penalized (about 33% slow down) by the larger number of propagator executions as well as domain updates, both again one order of magnitude higher than in the view models. Within these, PVIEWS performs slightly better due to function inlining (in this case, there are fewer function calls than the previous example even if all function calls were inlined).

| Golomb |        | p       | t    | f       | u      | calls  | op     |
|--------|--------|---------|------|---------|--------|--------|--------|
| 10     | VARs   | 3.64E+5 | 0.05 | 1,703   | 2.2E+5 | 0      | 1.9E+6 |
|        | SViews | 3.72E+4 | 0.04 | 1,707   | 3.0E+4 | 2.3E+5 | 1.7E+6 |
|        | PViews | 3.72E+4 | 0.04 | 1,707   | 3.0E+4 | 0      | 1.7E+6 |
| 11     | VARs   | 2.06E+6 | 0.3  | 7,007   | 1.3E+6 | 0      | 1.1E+7 |
|        | SViews | 1.84E+5 | 0.22 | 7,063   | 1.5E+5 | 1.2E+6 | 1.1E+7 |
|        | PViews | 1.84E+5 | 0.21 | 7,063   | 1.5E+5 | 0      | 1.1E+7 |
| 12     | VARs   | 1.14E+8 | 17.7 | 283,156 | 6.8E+7 | 0      | 6.1E+8 |
|        | SViews | 9.61E+6 | 13.8 | 284,301 | 8.3E+6 | 5.9E+7 | 6.6E+8 |
|        | PViews | 9.61E+6 | 13.4 | 284,301 | 8.3E+6 | 0      | 6.6E+8 |

Table 8: Number of propagator executions (p), seconds (t), fails (f), domain updates (u), calls to view object methods (calls), and arithmetic operations (op) when solving three instances of the Golomb problem.

## 6 Conclusion and Future Work

In this paper we addressed the modeling of decomposable constraints and challenged the traditional view that such decomposition is best made through the use of auxiliary variables. We showed that the propagation of such decomposable constraints may be performed by view based propagators that do not require such auxiliary variables, and discussed the properties of several approximations of this approach. In particular, we focused on models using box view-based propagators for pragmatic reasons, and showed that, notwithstanding the fact that asymptotic worst-case analysis leads to the same complexity, when applied to a comprehensive set of benchmarks they perform significantly better than those based on auxiliary variables, even when the latter models use stronger propagators.

We have intentionally restricted the instantiation of view models to those consisting only of box approximations, largely because box view objects may be implemented efficiently. In fact, other view models may be propagated using a similar approach, but creating efficient corresponding view objects is much more challenging in general. As an example, consider designing a domain view object, that is an object which propagates constraints of the form  $c \circ f_i \cdots \circ f_m$  and computes a  $\delta\delta$ -approximation for every image and object function of each of the functions  $f_i$ . Such object must thus be able to compute  $\delta$ -domains, which cannot be done in constant time for most functions  $f_i$  as it is the case for  $\beta$ -domains. But it will be interesting to check whether this could be done for specific classes of functions, similarly to what was done in [26].

Finally, given the superior performance of PVIEWES compared with SVIEWS, one may question what to do when the former cannot be used directly, as occurs when the problem at hand is specified in an interpreted environment such as MiniZinc [21] where there is no C++ compilation involved. In this case, an option is to generate C++ code for the constraints and compile it with a JIT (Just In Time) C++ compiler. Pre-processing constraint problems for generating specific solver binaries is shown in [1]. It would be interesting to see if the compilation time would compensate the solving time in our case.

## References

- [1] D. Balasubramaniam, C. Jefferson, L. Kotthoff, I. Miguel, and P. Nightingale. An automated approach to generating efficient constraint solvers. In *Proceedings of the 2012 International Conference on Software Engineering, ICSE 2012*, pages 661–671, Piscataway, NJ, USA, 2012. IEEE Press.
- [2] T. Ball, A. Podelski, and S. K. Rajamani. Boolean and Cartesian abstraction for model checking C programs. *International Journal on Software Tools for Technology Transfer (STTT)*, 5(1):49–58, 2003.
- [3] N. Beldiceanu, M. Carlsson, and J.-X. Rampon. Global constraint catalog. The Internet, 2010.
- [4] F. Benhamou. Interval constraint logic programming. In A. Podelski, editor, *Constraint programming: basics and trends*, volume 910 of *Lecture Notes in Computer Science*, pages 1–21. Springer, 1995.
- [5] F. Benhamou. Heterogeneous constraint solving. In *Algebraic and Logic Programming, ALP’96*, volume 1139 of *Lecture Notes in Computer Science*, pages 62–76. Springer-Verlag, 1996.
- [6] B. Carlson. *Compiling and executing finite domain constraints*. PhD thesis, Uppsala University, 1995.
- [7] K. C. K. Cheng and R. H. C. Yap. Maintaining generalized arc consistency on ad hoc r-ary constraints. In P. J. Stuckey, editor, *Principles and Practice of Constraint Programming, CP’08*, volume 5202 of *Lecture Notes in Computer Science*, pages 509–523. Springer, Sept. 2008.
- [8] M. Correia. *Modern Techniques for Constraint Solving: The CaSPER Experience*. PhD thesis, Universidade Nova de Lisboa, December 2010. Pedro Barahona (superv.);.
- [9] M. Correia and P. Barahona. On the Efficiency of Impact Based Heuristics. In *Principles and Practice of Constraint Programming, CP’08*, volume 5202 of *Lecture Notes in Computer Science*, pages 608–612. Springer, 2008.
- [10] M. Correia, P. Barahona, and F. Azevedo. CaSPER: A programming environment for development and integration of constraint solvers. In F. Azevedo, editor, *Workshop on Constraint Programming Beyond Finite Integer Domains, BeyondFD’05*, 2005.
- [11] A. M. Frisch, M. Grum, C. Jefferson, B. M. Hernandez, and I. Miguel. The essence of essence: A constraint language for specifying combinatorial problems. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, pages 73–88, 2005.

- [12] Gecode. Gecode: Generic constraint development environment. The Internet, 2010.
- [13] I. P. Gent, C. Jefferson, and I. Miguel. Minion: A fast scalable constraint solver. In *European Conference on Artificial Intelligence, ECAI'06*, pages 98–102. IOS Press, 2006.
- [14] I. P. Gent, C. Jefferson, and I. Miguel. Watched literals for constraint propagation in minion. In F. Benhamou, editor, *Principles and Practice of Constraint Programming, CP'06*, volume 4204 of *Lecture Notes in Computer Science*, pages 182–197. Springer, 2006.
- [15] I. P. Gent, C. Jefferson, I. Miguel, and P. Nightingale. Data structures for generalised arc consistency for extensional constraints. In *Conference on Artificial Intelligence, AAAI'07*, pages 191–197. AAAI Press, 2007.
- [16] I. P. Gent and T. Walsh. CSPLib: a benchmark library for constraints. Technical report, APES-09-1999, 1999.
- [17] P. V. Hentenryck, V. Saraswat, and Y. Deville. Constraint processing in cc(FD). Technical report, Brown University, 1992.
- [18] ILOG. *ILOG Solver 6.0 Reference Manual*. ILOG s.a., 2003.
- [19] R. Lopez-Ortiz, C.-G. Quimper, J. Tromp, and P. V. Beek. A fast and simple algorithm for bounds consistency of the alldifferent constraint. In *International Joint Conference on Artificial Intelligence, IJCAI'03*, pages 245–250. Morgan Kaufmann Publishers, Inc, 2003.
- [20] M. J. Maher. Propagation completeness of reactive constraints. In *International Conference on Logic Programming, ICLP'02*, volume 2401 of *Lecture Notes in Computer Science*, pages 148–162. Springer, 2002.
- [21] N. Nethercote, P. J. Stuckey, R. Becket, S. Brand, G. J. Duck, and G. Tack. Minizinc: Towards a standard CP modelling language. In C. Bessière, editor, *Principles and Practice of Constraint Programming, CP'07*, volume 4741 of *Lecture Notes in Computer Science*, pages 529–543. Springer, Sept. 2007.
- [22] F. Rossi, P. V. Beek, and T. Walsh, editors. *Handbook of Constraint Programming*. Foundations of Artificial Intelligence. Elsevier Science, 2006.
- [23] C. Schulte and P. J. Stuckey. Speeding up constraint propagation. In *Principles and Practice of Constraint Programming, CP'04*, volume 3258 of *Lecture Notes on Computer Science*, pages 619–633. Springer, 2004.
- [24] C. Schulte and G. Tack. Views and iterators for generic constraint implementations. In *Recent Advances in Constraints*, volume 3978 of *Lecture Notes in Computer Science*, pages 118–132. Springer, 2005.

- [25] C. Schulte and G. Tack. View-based propagator derivation. *Constraints*, pages 1–33, 2012.
- [26] G. Tack. *Constraint Propagation – Models, Techniques, Implementation*. Doctoral dissertation, Saarland University, Jan. 2009.